Cancer Causes and Control **14**: 645–655, 2003.

Prospective study of serum retinol, β-carotene, β-cryptoxanthin, and lutein/zeaxanthin and esophageal and gastric cancers in China

Christian C. Abnet[1,*], You-Lin Qiao[2], Sanford M. Dawsey[1], Dennis W. Buckman[3], Chung S. Yang[4], William J. Blot[5,6], Zhi-Wei Dong[2], Philip R. Taylor[1] & Steven D. Mark[7]
[1]*Cancer Prevention Studies Branch, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA;* [2]*Department of Epidemiology, Cancer Institute, Chinese Academy of Medical Sciences, Beijing, People's Republic of China;* [3]*Information Management Services, Rockville, MD;* [4]*Laboratory for Cancer Research, College of Pharmacy, Rutgers, The State University of New Jersey, Piscataway, NJ;* [5]*International Epidemiology Institute, Rockville, MD;* [6]*Department of Medicine, Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN;* [7]*Biostatistics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD, USA*

Received 31 December 2002; accepted in revised form 5 April 2003

Key words: β-carotene, β-cryptoxanthin, esophageal cancer, gastric cancer, lutein, retinol, zeaxanthin.

Abstract

Objective: This study examined the relationship between pretrial serum concentrations of retinol, β-carotene, β-cryptoxanthin, and lutein/zeaxanthin and the subsequent risk of developing esophageal squamous cell carcinoma and gastric cardia or non-cardia adenocarcinoma in subjects selected from a randomized nutritional intervention trial in Linxian, China, a region with epidemic rates of esophageal and gastric cardia cancer.
Methods: We used a stratified case–cohort design to select cohort members for inclusion in this study. In all we measured serum concentrations of the above vitamins in 590 esophageal, 395 gastric cardia, and 87 gastric non-cardia case subjects as well as in 1053 control subjects. Relative risks (RRs) were estimated using Cox proportional hazards models.
Results: Median values in our cohort were low for serum retinol (33.6 μg/dl), β-carotene (4.3 μg/dl), and β-cryptoxanthin (3.5 μg/dl) , but were high for lutein/zeaxanthin (40.0 μg/dl). Gastric cardia cancer incidence fell 10% for each quartile increase in serum retinol (RR = 0.90, 95% CI = 0.83–0.99). For esophageal cancer, an inverse association with retinol levels was found only in male non-smokers (RR = 0.79 per quartile increase, 95% CI = 0.63–0.99). For gastric non-cardia cancer, an inverse association was limited to subjects 50 years old or younger (RR = 0.58 per quartile, 95% CI = 0.31–0.96). For β-cryptoxanthin there was a borderline significant protective association for gastric non-cardia cancer (RR = 0.88 per quartile, 95% CI = 0.76–1.0). In contrast, we found the incidence of gastric non-cardia cancer increased (RR = 1.2 per quartile, 95% CI = 1.0–1.3) with increasing concentration of serum lutein/zeaxanthin.
Conclusions: In this population, we found that low retinol and high lutein/zeaxanthin concentrations increased the risks of gastric cardia and gastric non-cardia cancer respectively. We found that there were no strong associations between any of the other analytes and any of the cancer sites.

Abbreviations: CI – confidence interval; l – liters; μg – micrograms; ml – milliliters; QC – quality control; RR – relative risk; SD – standard deviation; y – years

* Address correspondence to: Christian Abnet, PhD, MPH, Cancer Prevention Studies Branch, National Cancer Institute, 6116 Executive Blvd, Room 705, Bethesda, MD 20892-7058, USA. Ph.: +1-301-594-1511; Fax: +1-301-435-8645; E-mail: abnetc@mail.nih.gov

Introduction

Antioxidant vitamins have long been thought to play a role in reducing the risk of cancers at many sites. In

particular, it has been hypothesized that the antioxidant activity of vitamins may be protective for gastrointestinal cancer. Antioxidants inhibit DNA synthesis, reduce cell proliferation, alter gene expression, inhibit cell transformation, and protect against oxidative stress (reviewed by Collins [1]).

Only two recent prospective studies have used serum measurements to assess the association of retinol and β-carotene with subsequent development of gastric cancer. A nested case–control study in Hawaii of individuals of Japanese descent found approximately 70% fewer incident gastric cancers in those in the highest tertile of serum β-carotene [2]. No effect of higher serum retinol was found. A recent study from an area of China with high gastric cancer rates reported no association between serum β-carotene or retinol and progression or regression of pre-neoplastic gastric lesions [3]. There have been no prospective studies and no serum studies of any design relating retinol, or carotenoids, to esophageal cancer risk.

Two randomized intervention trials, both conducted by our group in Linxian, China, have assessed the effects of supplementation with antioxidant vitamins and minerals on esophageal and gastric cancer [4–6]. The residents of Linxian had diets that were marginal in numerous nutrients, and high rates of esophageal squamous cell carcinoma and gastric cardia adenocarcinoma of more than 400 per 100,000 person-years, approximately 100 times the rates of US whites [7]. In the General Population Trial, 29,584 participants were sampled from the general population of Linxian. This trial tested four different combinations of nutrient supplements for 5.25 years. The group supplemented with selenium, β-carotene, and vitamin E (factor D) had a statistically significant reduction of 9% in all-cause mortality and 13% in cancer mortality [6]. The mortality and incidence rates for esophageal and gastric cardia cancers combined were reduced by 10 and 6% respectively [6]. The subjects in the General Population Trial comprise the cohort of individuals in our current study.

The primary goal of this study was to evaluate the relationship between baseline serum concentrations of retinol and the carotenoids β-carotene, β-cryptoxanthin, and lutein/zeaxanthin with the subsequent risk of esophageal squamous cell carcinoma, gastric cardia adenocarcinoma, or gastric non-cardia adenocarcinoma cases. In addition, we examined the modifying effects of age, sex, tobacco use, alcohol use, and intervention treatment group assignment on these associations. In studies of identical design we have analyzed the association of these cancers with serum selenium concentrations [8] and with serum α- and γ-tocopherol concentrations (P.R. Taylor *et al.*, in press, JNCI).

Subjects and methods

Cohort population

The subjects in this study were selected from the cohort of all participants in the General Population Trial of Linxian. Elsewhere we have described in detail the design, choice of intervention agents, methods of conduct, and primary end-point analyses of the trial [4, 6]. In brief, the participants were 29,584 healthy adults aged 40–69 years from four Linxian communes. In the spring of 1985, one year before the start of the intervention, each participant was interviewed, was given a brief physical examination, and had 10 ml of blood drawn. Intervention began in March 1986 and continued through May 1991. In accord with a partial factorial design, the participants were randomly assigned to receive either a vitamin–mineral combination or a placebo. In total, four different vitamin–mineral combinations were tested: factor A (10,000 IU of vitamin A and 45 mg of zinc oxide), factor B (52 mg of riboflavin and 40 mg of niacin), factor C (180 mg of ascorbic acid and 30 g of molybdenum), and factor D (50 μg of yeast, 15 mg of β-carotene, and 30 mg of α-tocopherol). Village doctors ascertained mortality among trial participants through monthly follow-up. Diagnoses of cancer were ascertained through local commune and county hospitals, and were supplemented by a study team that provided clinical and diagnostic services, including endoscopy, for patients with symptoms suggestive of esophageal or gastric cancer. A panel of US and Chinese experts reviewed the diagnostic material for 90% of the cancer cases in this study. For anatomic localization of gastric adenocarcinomas, we defined cardia cancers as those located in the most proximal three cm of the stomach; non-cardia cancers were those that originated outside this region. Ninety-five percent of the anatomic localizations were made with the use of endoscopy, surgery, and/or X-rays. For cancer cases without diagnostic material and for deaths due to causes other than cancer, senior Chinese diagnosticians conducted reviews. At the end of trial in May1991, interviews and brief physical examinations were conducted. We obtained written informed consent from each participant before trial enrollment. Throughout the trial, human subject protection procedures were followed in accord with those prescribed by the US National Institutes of Health and the Chinese Academy of Medical Sciences.

Selection of study participants for serum measurements

We used a stratified case–cohort design [9–11] to select individuals for serum measurements from the cohort of

all participants in the General Population Trial. By the end of the trial, there were 640 incident esophageal squamous cell carcinomas, 435 incident gastric cardia adenocarcinomas, and 104 gastric non-cardia adenocarcinomas [6]. Overall, 91% of the case subjects had adequate serum for anti-oxidant measurements (590 esophageal, 395 gastric cardia, and 87 gastric non-cardia cancer subjects, see Table 1). In addition, we measured serum concentrations in a stratified random sample of all trial participants, without regard to outcome, as the comparison group (referred to as the subcohort). The six strata were defined by sex and the following three age categories (age defined as age at start of the intervention): (1) 50 years old or younger, (2) older than 50–60 years old, and (3) older than 60 years. A sufficient number of cohort subjects were drawn from each stratum to achieve a ratio greater than 1:1 of control subjects to case subjects for incident esophageal and gastric cardia cancers combined. The lowest within-strata ratios of control subjects to case subjects for the incident site-specific cancers ranged from 1.4 to 2.0. Overall, we measured serum vitamin concentrations in 1053 subcohort subjects and 1072 case subjects.

Laboratory analysis

At the time of the baseline interview, March–May 1985 (one year prior to the initiation of the intervention), a 10 ml venous blood sample was collected from all consenting participants. Blood was stored 3–6 h on ice during transportation to the central field station where it was separated, aliquoted, and stored at −45 °C for 3–4 days prior to shipment to Beijing, where long term storage was at −85 °C. In 1996 collected sera were transferred to the National Cancer Institute repository on dry ice. Samples were thawed, additional aliquots made, and immediately refrozen prior to shipment on dry ice to the analytic laboratories. A modified simultaneous isocractic-HPLC assay based on a previously described method [12, 13] was used to determine serum retinol, β-carotene, β-cryptoxanthin, and lutein/zeaxanthin concentrations in the laboratory of Dr C.S. Yang in 1997. This method maximized the number of analytes that could be quantified with a small serum sample but did not allow for independent measurement of serum lutein and zeaxanthin concentrations. Cholesterol was measured by the NHANES laboratory, Centers for Disease Control and Prevention (Atlanta, GA). Serum

Table 1. Number, mean age[a], smoking, alcohol, and BMI[a] by case status in the nested case–cohort from the General Population Trial cohort, Linxian, China: overall and by sex

	Subcohort	Case subjects[b]		
		Esophagus	Gastric cardia	Gastric non-cardia
Total				
Number	1053	590	395	87
Mean age (SD)	56.4 (8.0)	56.4 (8.0)	57.1 (7.3)	58.3 (7.4)
Smoking %[c]	37.6	37.5	42.3	54.0
Alcohol %[d]	20.3	22.0	21.5	26.4
BMI (SD), kg/m^2	22.1 (2.5)	21.5 (2.4)	21.6 (2.5)	21.1 (1.8)
Females				
Number	479	304	162	21
Mean age (SD)	55.0 (8.4)	55.1 (8.3)	56.3 (7.4)	54.3 (7.9)
Smoking %	0.4	0.00	0.00	0.00
Alcohol %	6.1	6.3	9.3	4.8
BMI (SD), kg/m^2	22.3 (2.8)	21.4 (2.6)	22.1 (3.1)	21.5 (1.7)
Males				
Number	574	286	233	66
Mean age (SD)	57.5 (7.5)	57.7 (7.4)	57.7 (7.2)	59.7 (6.7)
Smoking %	68.6	77.3	71.7	71.2
Alcohol %	32.2	38.8	30.0	33.3
BMI (SD), kg/m^2	21.8 (2.1)	21.5 (2.0)	21.3 (2.0)	21.0 (1.7)

[a] Means and SD were calculated using the sampling weights from each of the six age/sex strata for each category.
[b] All esophageal cancers are squamous cell carcinomas and all gastric cancers are adenocarcinomas.
[c] Smoking defined as ever smoking ≥6 months.
[d] Alcohol use defined as any ethanol consumption in the previous 12 months.

specimens were analyzed on the Kodak Ektachem 250 Dry Chemistry Analyzer, using a single-slide two-point enzymatic cholesterol test, as recommended by the manufacturer (Eastman Kodak Co., Rochester, NY).

For β-carotene and β-cryptoxanthin 7.5 and 2.4% of the samples had values below the lower limit of detection (57 and 35 μg/dl respectively). For this analysis, we assigned these individuals serum values equal to one-half the limit of detection. Assigning values of zero or a value equal to the limit of detection did not alter the results. All measurements of lutein/zeaxanthin and retinol were above the lower limits of detection (22 and 27 μg/dl respectively). In addition to the above analytes, the serum was assayed for α-carotene and lycopene. Because 52% of subjects had α-carotene concentrations below the limit of detection (48 μg/dl), and 97% of subjects had lycopene concentrations below the limit of detection (56 μg/dl), no further analyses were conducted on these two analytes. This population consumes few if any tomatoes so the low serum concentrations of lycopene are expected. Previous measurements of α-carotene in this population demonstrated a mean serum concentration <3 μg/dl [14]. All of the cholesterol concentrations were above the limit of detection.

The samples were shipped and analyzed in a sequence designed to minimize the possible bias in the estimation of cancer risk that could be introduced if serum measurements varied by time. Within every group of 10 samples, a sample from a case subject was always accompanied by a sample from a control subject from the same sex and age stratum. Case subjects with each of the three cancer types in each of the six strata were mixed throughout. For assessment of assay reliability, approximately one of every 10 samples was a quality-control sample of serum obtained from processing pooled whole blood drawn in 1996 from three residents of Linxian. Laboratory personnel were unaware of the case status of samples and of the existence of these quality-control samples. In all, a total of 245 blinded quality-control specimens were measured. Coefficients of variation were as follows: retinol = 6.0, β-carotene = 17.6, β-cryptoxanthin = 10.6, lutein/zeaxanthin = 6.8, and cholesterol = 2.1.

To detect systematic variation of the measurement over time the quality-control specimens were plotted *versus* time, and a loess smooth [15] fit to the quality-controls. The optimum loess bandwidth for each analyte was chosen by minimizing the Cp statistic [15]. To assess whether the loess regression using the chosen bandwidth resulted in a statistically significant improvement we used an *F*-test that compared the residual sums of squares around the mean of the quality-control values with the residual sums of squares around the loess fit [16]. The retinol, β-carotene, β-cryptoxanthin, and lutein/zeaxanthin measurements all showed statically significant variation over time (all p's $< 10^{-5}$). For these analytes we estimated the time dependent measurement error as the difference between the median of the log quality-control value and the predicted loess quality-control at that time. We corrected the analyte values of the subjects by subtracting this estimate of measurement error from the subjects observed analyte value. The cholesterol measurement showed no significant variation over time and no corrections were made.

Statistical analysis

To graphically examine the shape of the serum distributions in the General Population Trial cohort, we used histograms, quantile–quantile plots, and plots of residuals from regressions of the analytes on sex, age, and cholesterol. For all analytes we found that log transformation improved normality. On the log scale there were no outliers or observations with high influence. All estimates of means and the difference between means were made using the log transformed values. In the Tables we report means and regression coefficients returned to the original scale. Throughout the paper, all p-values we report are two-sided.

The mean values and quantiles were calculated using the known sampling weights from the entire General Population Trial cohort for each individual in the study [17]. Thus, for example, the means and quantiles for the vitamins and cholesterol given in Table 2 are estimates of the means and quantiles of the entire General Population Trial cohort and not, as is generally the case, the means and quantiles of those who never develop the cancers under study. We examined the association among serum analytes and between the serum analytes and sex, age, cholesterol, smoking, drinking, and BMI using weighted Pearson correlation and weighted linear regression.

We measured the time to cancer as time since March 1986 (start of the intervention), as opposed to time since blood collection. Individuals who died or developed cancer in this one-year interval were excluded from the General Population Trial and this analysis. When analyzing cancers at a specific site we treated persons with cancers at other sites as censored at the time of cancer occurrence. We estimated relative risks (RRs) and 95% confidence intervals (CI) using the case–cohort estimator for the Cox proportional hazards models [9–11, 18]. All estimates came from models stratified on the six sex-age sampling strata. Additional stratum-specific age terms for continuous age were used to adjust for

Table 2. Means[a] and selected quantiles for the baseline vitamin and cholesterol concentrations in the General Population Trial cohort, Linxian, China: overall and by sex

	Number	Analyte	Analyte level					
			Geometric mean[b]	10%	25%	50%	75%	90%
Overall	2125	β-Carotene (μg/dl)	5.8	1.4	2.4	4.3	7.3	12.4
		β-Cryptoxanthin (μg/dl)	5.5	1.2	2.0	3.5	6.6	12.0
		Lutein/Zeaxanthin (μg/dl)	42.6	19.4	27.8	40.0	52.6	67.8
		Retinol (μg/dl)	34.9	19.6	25.3	33.6	42.1	50.8
		Cholesterol (mg/dl)	151.2	110.0	126.4	146.5	171.6	195.9
Females	966	β-Carotene (μg/dl)	7.2	2.0	3.3	5.3	9.4	14.1
		β-Cryptoxanthin (μg/dl)	6.6	1.6	2.6	4.2	7.9	14.2
		Lutein/Zeaxanthin (μg/dl)	46.9	24.1	32.1	43.5	56.8	73.1
		Retinol (μg/dl)	34.7	19.9	25.6	33.6	42.1	50.0
		Cholesterol (mg/dl)	159.7	117.2	133.3	153.5	182.5	203.9
Males	1159	β-Carotene (μg/dl)	4.2	0.4	1.8	3.1	5.2	8.0
		β-Cryptoxanthin (μg/dl)	4.0	0.9	1.6	2.6	4.7	9.4
		Lutein/Zeaxanthin (μg/dl)	37.3	16.4	23.2	34.4	47.7	60.3
		Retinol (μg/dl)	35.0	19.1	25.0	33.4	42.0	51.6
		Cholesterol (mg/dl)	140.6	105.0	118.6	137.3	158.3	174.8

[a] Means and quantiles were calculated using the sampling weights from each of the six age/sex strata. These estimate the distributions in the entire General Population Trial cohort of 29,584 subjects. The population quartile boundaries were used as the boundaries for all further analyses based on quartiles. When examined as continuous variables the values were centered and standardized to the average size of the two central quartiles.

[b] Mean serum concentrations for these analytes in the US NHANES III study for individuals 51–70 years were as follows: β-carotene 26.1 μg/dl, β-cryptoxanthin 10.3 μg/dl, lutein/zeaxanthin 24.1 μg/dl, retinol 61.3 μg/dl.

variation within age stratum. Nested models were compared using score tests. To examine the proportionality assumption, we used models that allowed time-dependent RRs. We found no suggestion that RRs varied with time. We performed a sensitivity analysis of all main-effect RRs by deleting the upper and lower one percent of serum values. There were no substantial changes in RR estimates.

For all RRs, we examined three different metrics for each analyte. Each vitamin was used as a continuous variable with the units standardized by the average size of a central quartile, half the difference between the values of quartile 3 and quartile 1. For each analyte the standardized units are as follows: Retinol 8.36 μg/dl, β-Carotene 2.45 μg/dl, β-Cryptoxanthin 2.30 μg/dl, and Lutein/Zeaxanthin 12.41 μg/dl. We also classified individuals into quartiles and estimated RR both by allowing a separate effect for each quartile and by assigning each individual the ordinal value of his/her quartile. In Table 3 we estimate the RRs for each analyte individually and report both the continuous and individual quartile RRs. The trend test p-value is generated from the ordinal model. In Table 4 we estimated the RR for each individual analyte while simultaneously controlling for the other analytes; in that table we only present RRs using the standardized analytes measured on the continuous scale.

We tested the Cox Proportional Hazards models for deviations from log linearity by adding quadratic terms to the continuous models and by adding separate parameters for individuals in the lowest and highest deciles (threshold effects). No statistically significant quadratic deviations or thresholds were found.

For each analyte we tested whether any of the cancer RRs varied by the covariates age, sex, tobacco use, alcohol use, or treatment factor. We did this by comparing a model with the main effect of a covariate (*e.g.* sex) and a single risk parameter for the analyte, to a model with the main effect term for the covariate and separate risk parameters for each sub-group (*e.g.* females and males) defined by the covariate. Since all models were stratified on sex and age, there was no main effect term when testing and estimating the sex interaction. To obtain estimates of the RR for each sub-group, models were restricted to individuals in a given sub-group. For example, within the categories of sex, males did not contribute to the RR estimate of the effect of analytes on cancer in females. Interactions between the analytes and smoking were limited to males since almost none of the females in our cohort smoke.

Table 3. RR[a] and 95% CI for changes in serum vitamin concentration and as quartiles by cancer site in the General Population Trial cohort, Linxian, China

	Continuous[c]			Quartiles[b]							
	RR	95% CI	*p*	1	2		3		4		p_{trend}[d]
				(Ref)	RR	95% CI	RR	95% CI	RR	95% CI	
β-Carotene											
Esophageal	1.0	0.95–1.1	0.94	1	1.1	0.82–1.5	0.89	0.66–1.2	1.0	0.74–1.4	0.72
Gastric cardia	1.0	0.96–1.1	0.52	1	0.89	0.64–1.2	0.76	0.54–1.1	0.95	0.67–1.4	0.54
Gastric non-cardia	1.0	0.91–1.1	0.76	1	1.9	1.0–3.6	1.3	0.66–2.6	1.9	0.89–3.9	0.25
β-Cryptoxanthin											
Esophageal	1.0	0.98–1.1	0.57	1	1.2	0.91–1.6	1.1	0.82–1.5	1.3	0.93–1.7	0.24
Gastric cardia	1.0	0.99–1.1	0.23	1	0.97	0.69–1.4	1.2	0.90–1.7	1.2	0.85–1.7	0.15
Gastric non-cardia	0.90	0.78–1.1	0.13	1	1.6	0.87–2.8	1.6	0.88–2.8	0.54	0.23–1.3	0.50
Lutein/Zeaxanthin											
Esophageal	0.96	0.90–1.1	0.24	1	1.0	0.80–1.4	0.80	0.60–1.1	0.89	0.66–1.2	0.22
Gastric cardia	1.0	0.96–1.1	0.49	1	1.0	0.75–1.4	0.76	0.54–1.1	1.1	0.81–1.6	0.87
Gastric non-cardia	1.1	1.0–1.3	0.035	1	2.0	1.1–3.8	2.2	1.2–4.2	1.7	0.82–3.6	0.10
Retinol											
Esophageal	0.97	0.91–1.0	0.43	1	1.0	0.77–1.4	1.1	0.80–1.4	0.97	0.71–1.3	0.96
Gastric cardia	0.95	0.87–1.0	0.18	1	0.82	0.60–1.1	0.55	0.39–0.78	0.82	0.59–1.2	0.067
Gastric non-cardia	1.0	0.86–1.2	0.97	1	0.70	0.38–1.3	1.2	0.67–2.2	0.79	0.39–1.6	0.94

[a] RR, 95% CI, and *p* values come from regression models stratified on sex and age with additional adjustment by separate continuous age variables for each age strata and variables for cholesterol, smoking, alcohol, and BMI.

[b] The RR, 95% CI for the continuous measure were standardized to the average size of the two central quartiles (see footnote for Table 2). Therefore this is the RR associated with a 25% change in serum concentration relative to the population distribution. One standardized unit is equal to: *β*-carotene 2.5 μg/dl, *β*-cryptoxanthin 2.3 μg/dl, lutein/zeaxanthin 12.4 μg/dl, and retinol 8.4 μg/dl.

[c] Quartiles were bounded as described in Table 2.

[d] The *p* for trend is derived from the score test for the addition of a variable where an individuals value is the quartile in which their serum value fell compared to a model with no variable for the specific analyte.

Table 4. RR[a] and 95% CI for changes in all four serum vitamin concentrations in a single model by cancer site in the General Population Trial cohort, Linxian, China

	β-Carotene		β-Cryptoxanthin		Lutein/Zeaxanthin		Retinol	
	RR[b]	95% CI	RR	95% CI	RR	95% CI	RR	95% CI
Esophageal	1.0	(0.96–1.1)	1.0	(0.98–1.1)	0.95	(0.88–1.0)	0.97	(0.90–1.1)
Gastric cardia	1.0	(0.96–1.1)	1.0	(0.99–1.1)	1.0	(0.95–1.1)	0.90	(0.83–0.99)
Gastric non-cardia	0.98	(0.86–1.1)	0.88	(0.76–1.0)	1.2	(1.0–1.3)	1.0	(0.85–1.2)

[a] RR, 95% CI, and p values come from regression models stratified on sex and age with additional adjustment by separate continuous age variables for each strata and variables for cholesterol, smoking, alcohol, and BMI.

[b] The RR, 95% CI for the continuous measure were standardized to the average size of the two central quartiles (see footnote for Table 2). Therefore this is the RR associated with a 25% change in serum concentration relative to the population distribution. One standardized unit is equal to: β-carotene 2.5 μg/dl, β-cryptoxanthin 2.3 μg/dl, lutein/zeaxanthin 12.4 μg/dl, and retinol 8.4 μg/dl.

Results

Table 1 presents the counts, the mean age, the mean BMI, and the percent smokers or drinkers by case status and sex. The incident cancers included 590 esophageal squamous cell carcinomas, 395 gastric cardia adenocarcinomas, and 87 gastric non-cardia adenocarcinomas. With the exception of patients with incident esophageal cancers, there were more male than female case subjects. For each of the six age and sex stratum, the difference between the mean age of all case subjects and the mean age of the subcohort subjects was not significantly different. We have previously reported the association between smoking, alcohol, and BMI and cancer incidence at these sites in Linxian [19]. Ever smoking was moderately associated, RR = 1.8, 95% CI = 1.4–2.4, with esophageal cancer but not stomach cancer. The generally moderate use of alcohol in this population was not associated with cancer at any of the three sites. BMI showed an inverse association with both esophageal and stomach cancer, where higher BMI is associated with reduced cancer risk. The risks for those in the highest quartile of BMI *versus* the lowest quartile were, RR = 0.7, 95% CI = 0.6–0.9 and RR = 0.8 95% CI = 0.6–1.0 for esophagus and total stomach cancer respectively [19]. There were significant sex differences in consumption of tobacco and alcohol. Both were dramatically lower in females compared to males. In this analytic cohort, 0% of the female cancer cases reported smoking and only 6.6% reported any alcohol consumption in the previous 12 months.

Using the known sampling weights, we estimated the geometric means and quantiles for each of the analytes and for cholesterol (Table 2) overall and by sex. These values estimate the population quantities for the entire General Population Trial cohort in 1985. The overall quantile values served as the boundaries for the models where we estimated cancer risk by quartiles.

We found significant ($p < 0.0001$) correlations between the serum concentrations of each the four vitamins and cholesterol. These ranged from from 0.25 to 0.39. Among the serum vitamins themselves, correlation coefficients ranged from 0.38 to 0.60, with the highest correlation occurring between β-carotene and lutein/zeaxanthin.

We examined whether the serum vitamin concentrations differ by sex, age, cholesterol, smoking, drinking, and BMI. Serum concentrations of all four of the analytes decreased with advancing chronological age. All three of the carotenoids had lower concentrations in males than in females. The largest difference occurred for β-carotene, where the mean concentration in males was only 60% as high as the mean concentration in females ($p < 0.0001$). All analytes were positively correlated with cholesterol. Smoking was not an important predictor for any of the analytes. Drinking and BMI were significant predictors for β-carotene only; drinkers had 10% higher β-carotene level than non-drinkers ($p = 0.040$) and β-carotene decreased by 6% for each quartile increase in BMI ($p < 0.0001$).

Tables 3 and 4 contain the RRs, 95% CIs, and tests of statistical significance relating vitamin concentrations to the subsequent development of site-specific cancers. The risk estimates in Table 3 are adjusted for age, sex, cholesterol, smoking, alcohol, and BMI but not for the other vitamins. Since the vitamins remained positively correlated even after adjusting for these other covariates, we assessed the independent effect of each nutrient using a model that simultaneously controlled for the effects of each of the other vitamins. These RRs are given in Table 4 using the continuous exposure scale.

One notable association in Table 4 was the decreased risk of gastric cardia cancer with increasing serum retinol concentration. Adjusting for the other vitamins, the risk of gastric cardia cancer decreased by 10% for each increase of 8.36 μg/dl retinol, (RR = 0.90, 95%

CI = 0.83–0.99). In the quartile analyses in Table 3, all the risks in quartiles 2, 3, and 4 were lower than in quartile 1. Although the test that this retinol-cancer association can be succinctly describe as a log linear trend is only borderline significant ($p = 0.07$), the overall global 3 df test of association showed that the risk of cancer in the upper three quartiles was significantly lower that in quartile 1 ($p = 0.01$).

Overall serum retinol concentration was not associated with esophageal cancers (Table 4) (RR = 0.97, 95% CI = 0.90–1.1). When individuals were classified by quartile of serum retinol, no consistent association was seen for incident esophageal cancers (Table 3). There was no association of retinol with gastric non-cardia cancer.

The other finding of note was a higher incidence of non-cardia cancers in those subjects with higher lutein/zeaxanthin concentrations. Table 4 shows that gastric non-cardia cancer incidence increased by 20%, (RR = 1.2, 95% CI = 1.04–1.3), for each 12.4 μg/dl increase in serum lutein/zeaxanthin concentration. The quartile analysis in Table 3 shows that individuals in quartile 2, 3 and 4 all have a similarly increased risk of approximately twofold when compared to quartile 1 individuals . The CIs for the risk estimates in quartiles 2 and 3 each exclude the null value of 1.

For β-carotene and β-cryptoxanthin we found no significant associations of vitamin concentration and cancer risk. Using the continuous scale, there was a borderline significant association for β-cryptoxanthin with gastric non-cardia cancer, (RR = 0.88, 95% CI = 0.76–1.0), but the quartile analysis showed no consistent pattern.

In estimating the cancer RRs in Table 3 we assumed that the effect of the vitamin concentrations on cancer rates was not modified by other individual characteristics. To evaluate this we examined the RR within groups defined by sex, age, smoking, drinking, randomization to factor A (10,000 IU of Vitamin A and 45 mg of zinc oxide), or factor D (50 μg of selenium, 15 mg β-carotene, and 30 mg α-tocopherol) separately. All effect modification analyses used the models described in Table 3 without correction for other serum vitamin concentrations. No higher order interactions were examined.

We found qualitatively important effect modification of the association of retinol and both esophageal ($p = 0.039$) and gastric non-cardia cancers ($p = 0.043$). For the esophagus, smoking was the effect modifier. In male non-smokers, higher retinol concentrations were associated with lower esophageal cancer rates, RR = 0.79, 95% CI = 0.63–0.99. In male smokers we found no such relationship , RR = 1.05, 95% CI = 0.94–1.17. Smoking was not an effect modifier of the association of retinol with either gastric cardia or non-cardia cancer. For gastric non-cardia cancer, effects differed by age category. For those in the youngest age stratum ($\leq$50), higher serum retinol concentrations were associated with lower risk of gastric non-cardia cancer, RR = 0.58, 95% CI = 0.31–0.96. In age strata 2 and 3 the RRs were not significantly different from 1, RR = 1.2, 95% CI = 0.97–1.4 and RR = 0.95, 95% CI = 0.76–1.2 respectively. No age interaction was seen for serum retinol concentration with either esophageal or gastric cardia cancers.

Age also modified the association of lutein/zeaxanthin with gastric non-cardia cancer ($p = 0.015$). The positive associations found in Tables 3 and 4 were confined to persons age 50 and older with RR = 1.2, 95% CI = 1.0–1.3 for subjects 50–60 years of age and RR = 1.1, 95% CI = 0.93–1.4 for subjects >60. No increased risk was seen for subjects <50 with RR = 0.91, 95% CI = 0.58–1.4. None of the RRs for the other serum vitamins had significant interactions with smoking or age.

One year after providing the serum on which these vitamin measurements were made, approximately half of the 2125 persons in this study began taking supplements containing retinol (factor A) and/or β-carotene (factor D). We found no evidence that treatment group modified the association of any of the baseline serum vitamin concentrations with subsequent cancer risk. In particular, the association of the of the pre-intervention serum retinol with subsequent cancer risks did not depend on whether an individual received retinol and zinc factor A). Similarly, receiving selenium, β-carotene, and vitamin E (factor D) did not alter any of the risk estimates for the associations of serum carotenoids with cancer incidence at any of the three sites.

No significant interactions were detected for sex, drinking or for any stratification of β-carotene or β-cryptoxanthin.

To address the concerns that occult preclinical cancers might have altered serum vitamin or cholesterol concentration and influenced RRs, we estimated the RRs separately for cases that occurred in the first two years of the intervention (one to three years after blood draw) and cases that occurred after the first two years. None of the estimates differed (data not shown). Similarly, allowing a time dependent RR in the proportional hazards models did not significantly improve the model fit.

Discussion

In this prospective study we examined the associations between baseline serum concentrations of retinol, β-

carotene, β-cryptoxanthin, and lutein/zeaxanthin, and the incidence of esophageal squamous cell carcinoma, gastric cardia adenocarcinoma and gastric non-cardia adenocarcinoma over 5.25 years. As expected, we found that serum concentrations of these vitamins differed from that of the US and Western populations. The median serum retinol concentration in our cohort (33.6 μg/dl) is near the 10th percentile in the NHANES III study [20]. Vitamin A deficiency is considered to occur when the serum retinol concentration is ≤20 μg/dl [21]. Thus approximately 12% of the Linxian population between the ages of 40–69 years were vitamin A deficient. This is similar to earlier reports from this population [14]. The median β-carotene concentration (4.3 μg/dl) is below the fifth percentile for all individuals examined in the NHANES III study [22]. For β-cryptoxanthin the median (3.5 μg/dl) also falls at approximately the fifth percentile for the NHANES III participants [22]. Ninety-seven percent of serum lycopene concentrations and 52% of serum α-carotene concentrations were below the limits of detection and therefore were excluded from further analysis. In contrast to the analytes described above, the median lutein/zeaxanthin concentration (40.0 μg/dl) is high, falling above the 95th percentile when compared to the distribution in the NHANES III [22]. The relatively low levels of fat soluble serum vitamins are not surprising because the residents of Linxian have other borderline nutritional deficiencies [14, 23], including B vitamins, which are also uniformly low compared to US residents (unpublished data). These low levels may limit the generalizability of our findings to replete populations.

In evaluating the association of serum vitamin concentrations with subsequent cancer risk we found that subjects with higher concentrations of serum retinol had lower incidence of gastric cardia cancer. We estimate that there was a 10% decrease in gastric cardia cancer incidence, RR = 0.90, 95% CI = 0.83–0.99, for each quartile increase of serum retinol.

We saw no significant overall association between serum retinol and esophageal cancer or non-cardia gastric cancers. However, in sub-groups defined by smoking status, we found that serum retinol concentrations were inversely associated with esophageal cancer risk in male non-smokers, RR = 0.79, 95% CI = 0.63–0.99, but not in male smokers, RR = 1.1, 95% CI = 0.94–1.2. The estimate for females RR = 0.96, 95% CI = 0.87–1.1, all of whom were non-smokers, was similar to that for the male smokers. We would have anticipated the retinol cancer relationship in male non-smokers to be similar to that in female cancer cases, all of whom were non-smokers. Of note, Linxian households are poorly ventilated and uncooked foods are contaminated with the potentially carcinogenic combustion products polycyclic aromatic hydrocarbons (PAH) [24, 25]. Thus females, who spend more time cooking, may be exposed to many inhaled carcinogens and resemble male smokers more than male non-smokers.

For gastric non-cardia cancers we also found an association of serum retinol and cancer incidence that was limited to a specific sub-group. For those in the youngest age stratum (≤50), higher serum retinol concentrations were associated with lower risk of gastric non-cardia cancer, RR = 0.58, 95% CI = 0.31–0.96. Those greater than 50 years of age showed no significant association between serum retinol and subsequent non-cardia cancer incidence. The relevance of this effect modification by age is difficult to assess. Inherent to sub-group analysis is the possibility that the multiple comparisons give rise to false positive findings. One potential explanation of this age group stratification in risk may be related to a birth cohort effect. The Red Flag Canal project, which was completed in the mid-1960s, improved the diet of the Linxian population through irrigated agriculture. It is conceivable that baseline serum measurements in the younger age group are more reflective of their typical adult serum vitamin exposure. Second, the higher cumulative exposure to environmental carcinogens, reflected in the higher absolute risk in older cohort members, may overwhelm the protective effect of retinol. The accuracy of these hypotheses cannot be directly assessed in our study design.

For lutein/zeaxanthin we found that for each quartile increase (12.4 μg/dl) there was 20% increase in gastric non-cardia cancer incidence (RR = 1.2, 95% CI = 1.0–1.3). A previous study of Chinese tin miners found that lung cancer risk increased with lutein/zeaxanthin concentrations [26]. In that population, as in the current study population, the serum lutein/zeaxanthin were high compared to those in Western countries. These elevations typify populations that consume large amounts of leafy green vegetables or corn. The high level of corn consumption in this population [27] suggests that the relatively high lutein/zeaxanthin is due to corn intake increasing zeaxanthin specifically. Corn consumption has been associated with esophageal cancer in other high-risk areas such as the Transkei region of South Africa [28]. Whether this reflects a poor diet with an over-reliance on grain, or a contaminant of corn, such as the mycotoxin fumonisin, is unknown. In another nested case–control study of the General Population Trial cohort, we saw no association between serum concentrations of sphingolipids, biomarkers of fumonisin exposure, and esophageal cancer, but that study did not examine gastric non-cardia cancer [29].

Two randomized trials reported that β-carotene supplementation increased the risk of lung cancer [30, 31]. Several hypotheses have been developed in an attempt to explain the mechanism of this effect. These hypotheses may be relevant to lutein/zeaxanthin if they are generalizable to other carotenoids. One mechanism proposes that chemicals that are antioxidants at low concentration become pro-oxidants at high concentrations [32]. A second hypothesis suggests that carotenoids may induce carcinogen activating enzymes. *In vitro* β-carotene has been shown to enhance the cell transforming activity of benzo[a]pyrene [33], presumably via induction of the phase I metabolizing enzymes CYP1A1 and CYP1A2 [34] and a subsequent increase in the activation of pro-carcinogens to reactive intermediates [33]. In our study, serum concentrations of β-carotene were low and had no association with gastric non-cardia cancer. In contrast, the concentrations of lutein/zeaxanthin were high in our population. If either the potential pro-oxidant effect or metabolic enzyme inducing effect of β-carotene is a general mechanism of carotenoids, the very high serum concentrations of lutein/zeaxanthin may have activated carcinogens in this PAH-exposed population [24].

This is the largest study to examine the association between retinol and carotenoids and risks of esophageal squamous cell carcinoma and gastric cardia adenocarcinoma. The large number of site-specific cancers, the close follow-up of the individuals in the cohort, and rigorous documentation of the cancer diagnoses, suggests that neither exposure nor disease misclassification are likely to have influenced our estimates. Though the effect of preclinical disease on serum measurements is always a concern, we found that risk estimates were not changed by considering time since blood draw.

There are several limitations which potentially reduce the power of our study to detect associations of vitamin concentrations and cancer risk. The follow-up time of 6.25 years after serum collection may fail to detect associations that might emerge with longer follow-up. Also, all serum studies of this design are limited by their reliance on a single measurement of vitamin concentration. Mitigating this short-coming is the relatively homogenous and monotonous diet of the Linxian population [27, 35] which should increase the ability for a single measure to characterize exposure. This homogenous diet also has the adverse effect of reducing the variation of vitamin concentrations in this population. For all the vitamins except lutein/zeaxanthin, the inter-quartile ranges in our cohort are considerably smaller than those in the NHANES III study.

Confounding by unmeasured risk factors is also a possible cause of associations in observational studies. The most important unmeasured risk factor for gastric cancer that we are aware of is *H. pylori* infection. We have previously demonstrated that *H. pylori* infection is a risk factor for both gastric cardia and non-cardia cancer in this population [36]. Other investigators have shown the *H. pylori* infection can alter both serum cholesterol concentrations [37] and gastric mucosa vitamin concentrations [37, 38]. Though we lack serologic measurement of *H. pylori* antibody for the majority of this cohort, we did have measurements for 352 subjects. Using this sub-group we estimated the correlation of the serum analytes with *H. pylori* positivity. All the correlation coefficients were small and none were statistically significant. In particular, *H. pylori* positivity and retinol were slightly negatively correlated ($r = -0.028$, $p = 0.59$), while *H. pylori* positivity and lutein/zeaxanthin were slightly positively correlated ($r = 0.062$, $p = 0.24$). The correlation with cholesterol was also small, but borderline statistically insignificant ($r = -0.096$, $p = 0.068$). The magnitude and direction of these associations and the known effect on cancer risk of *H. pylori* in this population makes confounding by *H. pylori* infection an unlikely explanation of our findings.

In summary, this is the largest prospective study to date that has examined the association between serum concentrations of retinol, β-carotene, β-cryptoxanthin, and lutein/zeaxanthin and incident esophageal squamous cell carcinoma and gastric cardia and non-cardia adenocarcinomas. In general, we found little association between the serum concentrations of these vitamins and cancer risk. For serum retinol, we found some evidence for all three cancer sites that higher serum concentrations are associated with lower cancer rates. The most compelling evidence is the association with gastric cardia cancer. For esophageal and gastric non-cardia cancer the association was found only within specific subgroups. Furthermore, we found that gastric non-cardia cancer occurred at higher rates in those with higher serum lutein/zeaxanthin concentrations.

References

1. Collins AR (2001) Carotenoids and genomic stability. *Mutat Res* **475**: 21–28.
2. Nomura AM, Stemmermann GN, Chyou PH (1995) Gastric cancer among the Japanese in Hawaii. *Jpn J Cancer Res* **86**: 916–923.
3. You W, Zhang L, Gail MH, *et al.* (2000) Gastric dysplasia and gastric cancer: *Helicobacter pylori*, serum vitamin C, and other risk factors. *J Natl Cancer Inst* **92**: 1607–1612.

4. Li B, Taylor PR, Li J-Y, *et al*. (1993) Linxian nutrition intervention trials. Design, methods, participant characteristics, and compliance. *Ann Epidemiol* **3**: 577–585.
5. Li JY, Taylor PR, Li B, *et al*. (1993) Nutrition intervention trials in Linxian, China: multiple vitamin/mineral supplementation, cancer incidence, and disease-specific mortality among adults with esophageal dysplasia. *J Natl Cancer Inst* **85**: 1492–1498.
6. Blot WJ, Li J-Y, Taylor PR, *et al*. (1993) Nutrition intervention trials in Linxian, China: supplementation with specific vitamin/mineral combinations, cancer incidence, and disease-specific mortality in the general population. *J Natl Cancer Inst* **85**: 1483–1492.
7. Blot WJ, Li JY (1985) Some considerations in the design of a nutrition intervention trial in Linxian, People's Republic of China. *Natl Cancer Inst Monogr* **69**: 29–34.
8. Mark SD, Qiao YL, Dawsey SM, *et al*. (2000) Prospective study of serum selenium levels and incident esophageal and gastric cancers. *J Natl Cancer Inst* **92**: 1753–1763.
9. Prentice RL (1986) A case cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* **73**: 1–11.
10. Self SG, Prentice RL (1988) Asymptotic distribution theory and efficiency results for case–cohort studies. *Ann Stat* **16**: 64–81.
11. Mark SD, Katki H (2001) Influence function based variance estimation and missing data issues in case–cohort studies. *Lifetime Data Anal* **7**: 331–344.
12. Miller KW, Yang CS (1985) An isocratic high-performance liquid chromatography method for the simultaneous analysis of plasma retinol, alpha-tocopherol, and various carotenoids. *Anal Biochem* **145**: 21–26.
13. Sowell AL, Huff DL, Yeager PR, Caudill SP, Gunter EW (1994) Retinol, alpha-tocopherol, lutein/zeaxanthin, beta-cryptoxanthin, lycopene, alpha-carotene, trans-beta-carotene, and four retinyl esters in serum determined simultaneously by reversed-phase HPLC with multiwavelength detection. *Clin Chem* **40**: 411–416.
14. Yang CS, Sun Y, Yang QU, *et al*. (1984) Vitamin A and other deficiencies in Linxian, a high esophageal cancer incidence area in northern China. *J Natl Cancer Inst* **73**: 1449–1453.
15. Hastie T, Tibshirani R. *Generalized Additive Models*. 1st edn. New York: Chapman and Hall, 1990.
16. *S-Plus 2000 Programmer's Guide*. Seattle, WA: MathSoft, Inc., 1999.
17. Shah BV, Barnwell BG, Bieler GS. *SUDAAN User's Manual*. Release 7.5 edn. Research Triangle Park: Research Triangle Institute, 1997.
18. *Epicure*. Seattle, WA: Hirosoft international Corp., 1998.
19. Guo W, Blot WJ, Li JY, *et al*. (1994) A nested case–control study of oesophageal and stomach cancers in the Linxian nutrition intervention trial. *Int J Epidemiol* **23**: 444–450.
20. Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc. Washington, DC: National Academy Press, 2000.
21. Sommer A, West KP Jr. *Vitamin A Deficiency: Health, Survival, and Vision*. New York: Oxford University Press, 1996.
22. Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids. Washington, DC: National Academy Press, 2000.
23. Ershow AG, Zheng SF, Li GY, *et al*. (1984) Compliance and nutritional status during feasibility study for an intervention trial in China. *J Natl Cancer Inst* **73**: 1477–1481.
24. Roth MJ, Strickland KL, Wang GQ, *et al*. (1998) High levels of carcinogenic polycyclic aromatic hydrocarbons present within food from Linxian, China may contribute to that region's high incidence of oesophageal cancer. *Eur J Cancer* **34**: 757–758.
25. Yang CS (1980) Research on esophageal cancer in China: a review. *Cancer Res* **40**: 2633–2644.
26. Ratnasinghe D, Forman MR, Tangrea JA, *et al*. (2000) Serum carotenoids are associated with increased lung cancer risk among alcohol drinkers, but not among non-drinkers in a cohort of tin miners. *Alcohol Alcohol* **35**: 355–360.
27. Yu Y, Taylor PR, Li JY, *et al*. (1993) Retrospective cohort study of risk-factors for esophageal cancer in Linxian, People's Republic of China. *Cancer Causes Control* **4**: 195–202.
28. Marasas WF, Jaskiewicz K, Venter FS, van Schalkwyk DJ (1988) Fusarium moniliforme contamination of maize in oesophageal cancer areas in Transkei. *S Afr Med J* **74**: 110–114.
29. Abnet CC, Borkowf CB, Qiao Y-L, *et al*. (2001) Sphingolipids as a Biomarker of Fumonisin Exposure and Risk of Esophageal Squamous Cell Carcinoma in China. *Cancer Causes Control* **12**: 821–828.
30. Albanes D, Heinonen OP, Taylor PR, *et al*. (1996) Alpha-Tocopherol and beta-carotene supplements and lung cancer incidence in the alpha-tocopherol, beta-carotene cancer prevention study: effects of base-line characteristics and study compliance. *J Natl Cancer Inst* **88**: 1560–1570.
31. Omenn GS, Goodman GE, Thornquist MD, *et al*. (1996) Risk factors for lung cancer and for intervention effects in CARET, the Beta-Carotene and Retinol Efficacy Trial. *J Natl Cancer Inst* **88**: 1550–1559.
32. Omaye ST, Krinsky NI, Kagan VE, *et al*. (1997) beta-carotene: friend or foe? *Fundam Appl Toxicol* **40**: 163–174.
33. Perocco P, Paolini M, Mazzullo M, Biagi GL, Cantelli-Forti G (1999) beta-carotene as enhancer of cell transforming activity of powerful carcinogens and cigarette-smoke condensate on BALB/c 3T3 cells in vitro. *Mutat Res* **440**: 83–90.
34. Paolini M, Cantelli-Forti G, Perocco P, *et al*. (1999) Co-carcinogenic effect of beta-carotene. *Nature* **398**: 760–761.
35. Li JY, Chen ZJ, Ershow AG, Blot WJ (1985) A case–control study of esophageal cancer in Linxian, People's Republic of China. *Natl Cancer Inst Monogr* **69**: 5–7.
36. Limburg PJ, Qiao Y-L, Mark SD, *et al*. (2001) Helicobacter pylori seropositivity and subsite-specific cancer risks in Linxian, China. *J Natl Cancer Inst* **93**: 226–233.
37. Hoffmeister A, Rothenbacher D, Bode G, *et al*. (2001) Current infection with Helicobacter pylori, but not seropositivity to Chlamydia pneumoniae or cytomegalovirus, is associated with an atherogenic, modified lipid profile. *Arterioscler Thromb Vasc Biol* **21**: 427–432.
38. Zhang ZW, Patchett SE, Perrett D, Domizio P, Farthing MJ (2000) Gastric alpha-tocopherol and beta-carotene concentrations in association with Helicobacter pylori infection. *Eur J Gastroenterol Hepatol* **12**: 497–503.